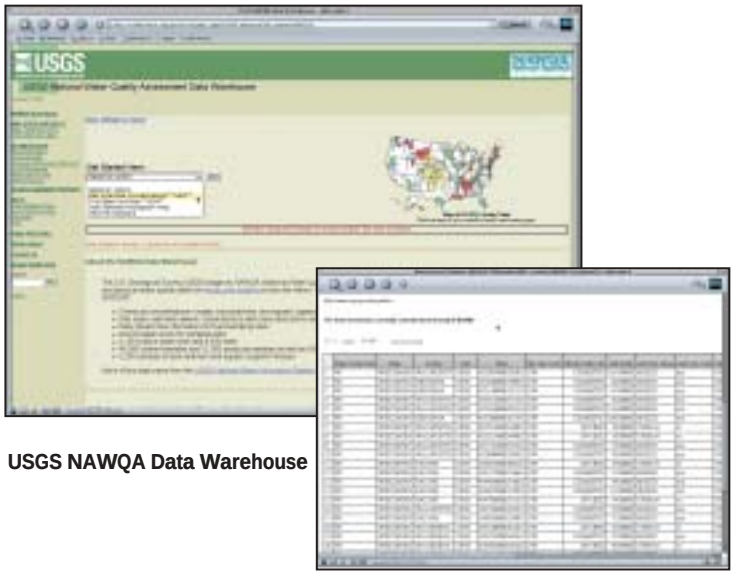
DATABASE APPLICATIONS DEVELOPMENT TEAM

Mission

Provide data organization and distribution services for water resources-related information at local, state, national, and international levels. Targeted organizations include local planning commissions, municipalities, the Wisconsin Department of Natural Resources, USGS, EPA, USAID, and others. An emphasis will be placed on using the Oracle RDBMS system and other Oracle tools for this purpose. Efforts will be made to acquire and develop solutions with the best state-of-the-art technologies available, including high-end data query tools, case tools, WWW-oriented distribution strategies, map serving software, and sophisticated relational database and multi-dimensional OLAP systems.



USGS NAWQA Data Warehouse

Team Members

Harry R. House, Hydraulic Engineer
John F. Walker, Research Hydrologist/Engineering
David W. Hall, Hydrologist
Daniel J. Sullivan, Hydrologist
Nathaniel L. Booth, IT Specialist (Systems Analysis)
Morgan A. Schneider, Hydrologist
Judith C. Thomas, Hydrologist
Carolyn J. McCullough, Hydrologist
James R. Statz, Information Technology Specialist

PROJECTS

DATABASE APPLICATIONS, WI 21000

COOPERATORS:

City of Milwaukee;
U.S. Environmental Protection Agency (USEPA);
U.S. Geological Survey, CAPP and WRD Wisconsin Department of Natural Resources (WDNR)

PROJECT CHIEF:

Harry R. House

LOCATION:

United States

PROJECT NUMBER:

WI 21000

PERIOD OF PROJECT:

July 1998–Continuing

PROBLEM

Natural resources agencies are having difficulty organizing, storing, and distributing their information products using their existing resources (staff, hardware, and software).

OBJECTIVE

The purpose of this project is to provide our cooperators, the USGS, and other agencies with modern alternatives to their existing information technology resources and capabilities. The objective is to provide our customers with access to high-quality data processing methods.

APPROACH

The Oracle Team in the Wisconsin District is dedicated to the design, development, and deployment of relational database systems geared to address water-resources-related data management problems being faced by our cooperators and other Federal agencies. Systems developed are centered on Oracle Corporation technology at this time. The team provides an alternative resource for these customers to compliment their in-house information technology staff and/or outside consultants.

The Oracle Team applies cutting-edge technology in the resolution of the data management problems, within the limitations of available funding and staff. It is assumed that lower-end technology or approaches are more widely available to our customers already, and that we will provide a more valuable service by providing higher-end systems. The team attempts to maximize the productivity of each developer by using development tools that reduce the need for custom coding, or by simply buying out-of-the-box applications that can be configured and used for agency needs.

PROGRESS (July 2001 to June 2002)

Eight new servers were purchased to support the increasing resource demands of the unit. Support licensing was renewed for all software, and additional purchases of software for web-enabled graphing was purchased from Visual Numerics Corporation. A

7-track tape backup system was purchased to further automate the backup procedures for the expanding database network. A full T 1 line was established with a local provider to improve bandwidth to users outside the USGS network. A 10-port 1000 Mb switch was added to improve network transfer speeds between internal team servers. An initial purchase of Redhat Linux was made to investigate the feasibility of putting Oracle middle tier and database tier software on existing Intel hardware, to improve stability, performance, and availability of systems to run newer versions of Oracle software. A large number of publicity articles and bulletins came out describing the teams activities with various IT products. A presentation was made at Oracle Open World describing the teams work on the NAWQA Data Warehouse.

The USGS Products Warehouse (subproject 21002) continued development, and is available for general use in the spring of 2002.

The Milwaukee Beach Health project (subproject 21003) with the City of Milwaukee was refunded, and is available for use by the public and internal users again.

A new geospatial-mapping interface was created for the NAWQA Data Warehouse (subproject 21004) in MapInfo java extreme. New data were entered into the system from the legacy sources.

The WDNR biological database (subproject 21005) added a module for tracking fish propagation information. Additional discoverer reports were created, and the existing forms were improved in response to user requests.

A visit to Russia to evaluate the potential for complimenting the team with selected Russian hydrologist/programmers was conducted in April 2002.

PLANS (July 2002 to June 2003)

It is expected that there will be additional purchases of servers to further enhance our hardware infrastructure. If an initial test of Redhat Linux is successful, there will likely be a whole-scale conversion to Linux on most servers in the summer/fall of 2002.

Is is anticipated that all current projects except 21003 will be refunded in the next period, and in some cases, at reduced levels in reaction to federal and state cutbacks. Project development activities are expected to make up the difference, if it occurs. In particular, a pilot for an NWIS internal data warehouse, and a Gateway to the Earth geospatial web search engine may prove to be the most likely candidates in this regard.

The productivity of the Russian collaboration will be evaluated, and if deemed successful, it is likely that program will be continued and/or expanded in the future.

Subproject 21002: we will continue gathering and loading citation data in cooperation with USGS Libraries. Internal testing and continued development of the web interface. Testing will begin with a small group and expand with each iteration. Comparisons will be made of completeness of data set with USGS Information Services data sets. We will be working on interfacing with USGS Information Services to provide ordering and pricing information for the products available. We plan to go into production sometime this summer with national publication citation information. In addition, we will create a multi-media database storage and deployment system associated with the data warehouse. The goal is to have the site act as the central repository for WRD for that type of information.

Subproject 21003: We will begin preparation and updating a site for 2002 swimming season. We expect to migrate the site into the newest version of Oracle Portal.

Subproject 21004: There are plans to develop a web-based graphing interface to produce constituent occurrence tables real-time. This would save NAWQA staff weeks of effort in producing these graphics for the 1997 Summary Reports. A re-working of aggregation plans to be done this period to accommodate new NWIS 4.1 structure. Also, there are plans to continue system documentation of all the NAWQA Data Warehouse data and application systems.

Subproject 21005: It is likely that substantial budget cuts at WDNR in FY 2003 will negatively influence the scope of project maintenance and development. While the FY 2003 funding problems have not improved since the last review period, there is much more satisfaction this review period with the performance of the Biology database among DNR management and staff since the T 1 line was installed and the new Fish Propagation Database was successfully rolled out. This new user satisfaction should at least build some support within DNR for the continued operation and funding of the Biology Database project during FY 2003.

Loading of the Ed Emmons/Don Fago dataset of fish and habitat data will be ongoing until completion.

Major revision of all the main database forms and reports is currently scheduled to consume 100% of the database development time for March through September 2002. Major changes include incorporation of a new Survey ID number into the database and may also involve converting most existing Designer Forms into more powerful and user-friendly Developer Forms. It has not yet been determined if construction of a new STAR schema for query support will be constructed as part of this major database revision.

WISCONSIN DEPARTMENT OF NATURAL RESOURCES BIOLOGY DATABASE

COOPERATOR:

Wisconsin Department of Natural Resources

PROJECT CHIEF:

David W. Hall

LOCATION:

Statewide

PROJECT NUMBER:

WI 21005

PERIOD OF PROJECT:

January 2000–Continuing

PROBLEM

The Wisconsin DNR Bureau of Fisheries Management and Habitat Protection collects and manages a wide variety of biological data. Data from biological programs were previously managed and stored in multiple databases in a wide variety of formats, making retrieval and analysis of data from the different programs difficult and time consuming. Databases did not exist and were needed to manage the statewide fish propagation program data and also for the statewide habitat assessment programs.

OBJECTIVE

A single database was needed to unify data from five legacy databases and two new databases. The legacy databases are the Wisconsin Inland Fisheries (Paradox) Database, Wisconsin Fisheries Historical Database, Treaty Database, Master-Fish File Database, and the Macroinvertebrate Database. The two new databases that were developed and were incorporated into the project include the Fish Propagation Database and the Habitat Database.

APPROACH

Initial development of the project was directed toward providing data entry and data reporting capability on the Internet for data from all major field activities. The most widely used field forms for fish, habitat, and macroinvertebrate data were selected for initial development. The database was deployed over the Internet so that geographically-dispersed users across the State could input and access data using electronic forms and reports.

PROGRESS (July 2001 to June 2002)

The URL for the database website is: http://infotrek.er.usgs.gov/pls/WDNR_BIOLOGY_WDB/WDNR_BIOLOGY_WDB.home. The database website, forms, and reports have undergone continual refinement in response to user needs in addition to ongoing development of new applications for the database. The first phase of the database was deployed on the Internet on January 1, 2001. Database use increased substantially and has remained high since November 2001 when the USGS Wisconsin District installed a

direct T-1 line to the Internet which substantially increased database performance. The new propagation component of the database went online in January 2002 and includes data-capture and reporting for processes from stocking-quota development, hatchery production, egg and fish transport, and final stocking. The planned major redesign and refinement of all database applications was undertaken from February through June 2002.

PLANS (July 2002 to June 2003)

The basic structure of the fish and habitat components of the database will be redesigned from June through September 2002 to link data from related fish and habitat surveys with a comprehensive survey ID. Additional data from legacy datasets will be loaded as time and funding allow. All database hardware and software are continually being upgraded to increase performance and to enhance the utility of the database applications. A feasibility study may be continued to assess the utility of mobile applications and electronic field data-entry forms as a paperless field data-collection system.